



Un modèle de l'apprenant constitué de plusieurs réseaux bayésiens concurrents.

Mathieu Hibou, Jean-Marc Labat

► To cite this version:

Mathieu Hibou, Jean-Marc Labat. Un modèle de l'apprenant constitué de plusieurs réseaux bayésiens concurrents.. Environnements Informatiques pour l'Apprentissage Humain (EIAH 2007), Jun 2007, Lausanne, Suisse. pp.413-424. hal-00161596

HAL Id: hal-00161596

<https://hal.science/hal-00161596>

Submitted on 11 Jul 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Un modèle de l'apprenant constitué de plusieurs réseaux bayésiens concurrents.

Mathieu Hibou*, Jean-Marc Labat**

* CRIP5-AIDA

Université René Descartes

45, rue des Saints-Pères

75270 Paris Cedex 06

mathieu.hibou@math-info.univ-paris5.fr

** LIP6-AIDA

Université Pierre et Marie Curie

12, rue Cuvier

75270 Paris cedex 05

jean-marc.labat@upmc.fr

RÉSUMÉ. Une des principales questions qui se pose en modélisation de l'apprenant à l'aide de réseaux bayésiens est celle de l'orientation des arcs. Notre hypothèse centrale est que cette orientation est liée au niveau d'expertise de l'apprenant, et que par conséquent un tel modèle devrait être composés de plusieurs réseaux différents. Nous donnons dans cet article une série d'arguments théoriques en faveur de cette hypothèse, puis nous présentons des modèles probabilistes constitués de plusieurs réseaux bayésiens concurrents (les multi-réseaux). Ces modèles nous permettent de mettre en œuvre nos idées et de tester nos hypothèses sur des données ; les résultats obtenus confortant notre approche.

MOTS-CLÉS : modélisation de l'apprenant, réseaux bayésiens, diagnostic cognitif.

1. Introduction

Les réseaux bayésiens sont d'efficaces outils de modélisation de l'apprenant dans l'incertain. Au cours de la dernière décennie, ils ont été utilisés avec succès dans de nombreux systèmes, avec différents objectifs, de l'évaluation des connaissances de l'apprenant à la reconnaissance du plan suivi en résolution de problème. De nombreux modèles ont ainsi été construits, et, à l'exception notable du travail de Reye [REYE 2004], il n'y a pas eu de tentative d'approche synthétique du problème. Notre objectif n'est pas ici de démontrer le bien fondé de l'utilisation de réseaux bayésiens en modélisation de l'apprenant (ils sont utilisés), mais d'étudier certaines modalités de leur mise en œuvre dans ce cadre.

Notre travail de thèse s'est centré sur la question de l'orientation des arcs, et de manière plus général sur la structuration des réseaux bayésiens en modélisation de l'apprenant. Nous tentons de montrer dans cet article en quoi cette question est cruciale. De plus, la question de l'adaptation structurelle du réseau aux comportements de l'apprenant à parfois été évoquée [MAYO&MITROVIC 2001], mais uniquement pour signaler l'existence du problème, sans apporter de piste ou d'ébauche de réponse, et ce alors que différents résultats de psychologie cognitive atteste de l'existence de différences structurelles selon le niveau d'expertise [CHI *et al* 1981]. L'hypothèse centrale de notre travail est qu'il existe un lien entre la structure du modèle de l'apprenant et son niveau d'expertise.

La problématique de cet article peut se résumer de la manière suivante : pourquoi et comment prendre en compte plusieurs réseaux bayésiens concurrents au sein d'un même modèle de l'apprenant ? Cette prise en compte est-elle expérimentalement justifiée ? Dans un premier temps, et après être revenu succinctement sur la définition des réseaux bayésiens, nous évoquons plus en détails les difficultés rencontrées lors de l'utilisation de réseaux bayésiens en modélisation de l'apprenant (en nous concentrant sur le cas le plus courant, c'est-à-dire celui de la construction par élicitation de savoir expert). Ensuite nous présentons les multi-réseaux qui sont des modèles graphiques probabilistes permettant de prendre en compte plusieurs réseaux au sein d'un même modèle et qui offrent par conséquent un cadre dans lequel tester nos hypothèses. La section 4 est quant à elle consacrée à la mise en œuvre d'un modèle de l'apprenant multi-réseau dans un cas concret et aux résultats expérimentaux obtenus.

2. Problèmes rencontrés lors de l'utilisation des réseaux bayésiens en modélisation de l'apprenant.

2.1. Quelques mots sur les réseaux bayésiens

Avant de préciser nos interrogations à propos de l'utilisation des réseaux bayésiens en modélisation de l'apprenant, il convient de revenir rapidement sur leur définition et sur la signification de l'inférence dans ce cadre.

2.1.1. Définition

Un réseau bayésien [PEARL 1988] est un graphe orienté sans circuit dont les nœuds sont des variables aléatoires et dont la topologie traduit les dépendances entre celles-ci. Plus formellement, un réseau bayésien est un couple (G, V) , G étant un graphe orienté sans circuit et V un ensemble de variables X_1, \dots, X_n en bijection avec l'ensemble des nœuds de G dont la loi jointe est donnée par

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | pa(X_i)) \text{ où } pa(Nœud) \text{ est l'ensemble des parents du nœud } Nœud \text{ dans } G.$$

Le graphe G est appelé *structure* du réseau, et l'on désigne par *paramètres* les tables de probabilités conditionnelles des nœuds.

2.1.2. Inférence et circulation de l'information

L'inférence dans un réseau bayésien est un calcul de probabilités conditionnelles : elle consiste en une mise à jour des probabilités après observation des valeurs d'un certain nombre de variables. En modélisation de l'apprenant, les variables observables modélisent les comportements de l'apprenant au travers de ses interactions avec le système. Ainsi, dans *Andes* [CONATI *et al* 2002], ce sont les capacités de l'apprenant en mécanique newtonienne qui sont inférées à partir de ses interactions lors de la résolution de problèmes. La circulation de l'information, c'est-à-dire l'influence qu'a l'observation d'une variable donnée sur les autres variables du graphe dépend de la structure du réseau. La figure 1 résume la circulation de l'information dans une chaîne selon qu'elle contient ou non un nœud observé (ou l'antécédent d'un nœud observé).

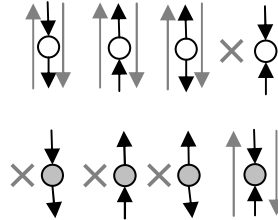


Figure 1. Circulation de l'information (les nœuds grisés sont observés ou sont les antécédents de nœuds observés).

Dans la suite de cet article, nous reprenons une typologie des nœuds inspirés de [CONATI *et al* 2002], et qui se retrouve sous différents vocables dans la littérature. La couche domaine est l'ensemble des nœuds modélisant les connaissances épistémiques sur l'apprenant, et la couche tâche celui des nœuds modélisant les actions de l'apprenant.

2.2. Orientation des arcs

La définition de la structure d'un modèle de l'apprenant basé sur un réseau bayésien à partir de savoir expert, se fait très souvent à partir d'une représentation

des connaissances du domaine (par exemple les connaissances en avionique dans [MISLEVY&GITOMER 1996] ou la programmation en Java dans [ZAPATA-RIVERA&GREER 2002]). Nous ne discutons pas ici de la validité du squelette (graphe non orienté) qui en découle en fonction du type de liens qui sont pris en compte (principalement des liens partitifs ou de généralité). Cependant, l'obtention de la structure du modèle à partir de ce squelette nécessite l'orientation des arcs. Ces orientations conditionnent le diagnostic obtenu, dans la mesure où elles jouent un rôle déterminant dans les relations de dépendance entre les variables.

2.2.1. *Influence de l'orientation des arcs sur le diagnostic*

Précisons sur un exemple l'influence de l'orientation des arcs sur le diagnostic obtenu. La figure 2 présente une version simplifiée du modèle de *Hydrive*, dans lequel les performances de l'apprenant en électronique sont observées. L'ensemble des variables représentant les différentes compétences de l'apprenant est affecté par cette observation. Ainsi, si l'apprenant se trouve être compétent en électronique, du fait par exemple de sa formation initiale, le modèle diagnostique qu'il en va de même dans l'ensemble des autres compétences.

Considérons (figure 3) un graphe ayant le même squelette, mais avec d'autres orientations des arcs. La propagation de l'information est ici bien plus limitée.

Non seulement le diagnostic obtenu dépend des choix d'orientation des arcs, mais il semble bien que l'orientation dépend du niveau d'expertise l'apprenant. En effet, s'il semble raisonnable de considérer le réseau de la figure 2 pour un sujet ayant suivi l'ensemble de la formation (pour lequel on peut supposer une certainement homogénéité des niveaux de compétence), le réseau 3 semble plus adéquat pour un sujet en début de formation, qui peut être très compétent dans un domaine particulier du fait de son cursus, sans pour autant maîtriser l'ensemble des compétences requises.

2.2.2. *Quels choix pour l'orientation des arcs ?*

Dans la littérature, le choix se porte très massivement en faveur d'une orientation des nœuds de la couche domaine vers ceux de la couche tâche.

La question reste ouverte en ce qui concerne les arcs reliant entre eux les nœuds de la couche domaine. Il est courant de présenter ce choix comme une alternative entre une orientation du général vers le particulier ou son contraire. On trouve dans la littérature des exemples de ces deux choix [MISLEVY&GITOMER 1996], [ZAPATA-RIVERA&GREER 2002]), sans que la justification apportée à ce choix soit toujours totalement convaincante, si tant est qu'elle existe. Cette dichotomie elle-même n'est à notre connaissance jamais discutée, la notion de réseaux bayésiens équivalents au sens de Markov [VERMA&PEARL 1991] permettant pourtant en partie de l'expliquer.

L'orientation des arcs du général vers le particulier convient lorsque les compétences de l'apprenant présentent une certaine homogénéité du fait des

dépendances que cette orientation induit. L'orientation dans le sens contraire est quant à elle plus adéquate pour des apprenants aux compétences hétérogènes. De plus, si nous tenons des constatations faites au §2.2.1 concernant le lien entre structure du réseau et niveau d'expertise de l'apprenant, nous concluons qu'il faut que le modèle permette cette évolution structurelle.

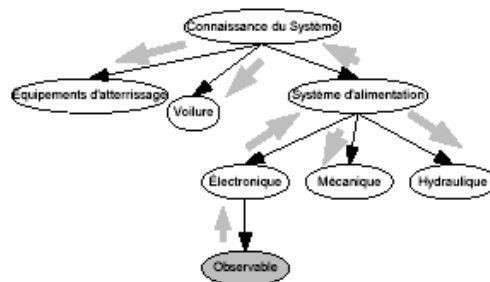


Figure 2. Flux d'information dans le modèle de l'apprenant de Hydrive.

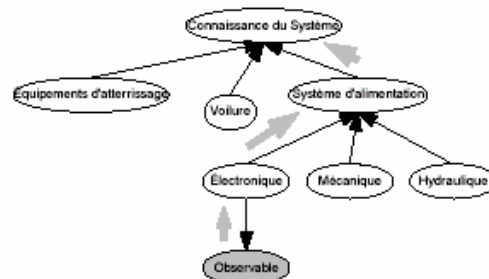


Figure 3. Flux d'information dans le modèle de l'apprenant de Hydrive modifié.

2.3. Asymétrie de dépendance

En probabilité on parle d'asymétrie de dépendance lorsque deux variables sont dépendantes étant données certaines valeurs de l'une d'entre elles et indépendantes étant données d'autres valeurs [GEIGER&HECKERMANN 1996].

Dans [MANSKE&CONATI 2005] est présenté un cas typique d'asymétrie de dépendance en modélisation de l'apprenant (qui n'est pas repéré comme tel)). Il s'agit de la description d'un modèle, constitué d'un réseau bayésien, représentant les connaissances de l'apprenant sur les décompositions des entiers en facteurs premiers. La construction du modèle de l'apprenant se fait à partir de la constatation suivante des enseignants : si un élève connaît la décomposition d'un nombre, alors (en général) il connaît celles de ses diviseurs, et le contraire est faux. En d'autres termes les connaissances de la décomposition d'un nombre et de celles de ses diviseurs ne sont dépendantes que dans le cas où l'élève connaît la décomposition du nombre en

question : il s'agit bien d'un cas de dépendance asymétrique (puisqu'elle dépend des valeurs des variables). Cette situation ne peut donc pas être modélisée par un réseau bayésien. Par conséquent, le modèle construit (qui est un réseau bayésien) ne peut pas réellement tenir compte des observations des enseignants.

2.4. Bilan

L'analyse de la littérature nous montre d'une part que la question de l'orientation des arcs est à la fois cruciale et en regard de quoi peu traitée, et d'autre part qu'un modèle devrait pouvoir prendre en compte plusieurs réseaux bayésiens que ce soit pour tenir compte de l'évolution structurelle de ses connaissances ou des relations de dépendances asymétriques.

3. Un cadre permettant la prise en compte de plusieurs réseaux au sein d'un même modèle : les multiréseaux.

Dans cette section, nous présentons les multi-réseaux qui sont des représentations probabilistes des connaissances permettant la prise en compte des asymétries de dépendances introduits dans [GEIGER&HECKERMANN 1996].

Les multi-réseaux offrent un cadre théorique de prise en compte de la variation de la structure des connaissances de l'apprenant : celui-ci n'est plus représenté par un unique réseau, mais par un ensemble de réseaux de structures différentes. Ce cadre théorique permet également de tester l'hypothèse centrale de notre travail sur le lien entre la structure des connaissances de l'apprenant et la structure du réseau qui le modélise (cf. §4).

3.1. Présentation

Les multi-réseaux généralisent les réseaux bayésiens dans le même sens que les mélanges de gaussiennes généralisent les distributions gaussiennes : un multi-réseau est constitué de plusieurs réseaux bayésiens à chacun desquels est associée une probabilité. Les différents réseaux bayésiens le constituant sont formés des mêmes noeuds, mais présentent des topologies différentes.

Du fait de son caractère généralisateur (un réseau bayésien est un multi-réseau à une seule composante), ce formalisme permet de tirer pleinement avantage de l'expertise acquise en modélisation de l'apprenant à l'aide de réseaux bayésiens.

Dans la suite, nous supposons que le multi-réseau est composé de n réseaux bayésiens r_1, \dots, r_n , dont les probabilités respectives sont notées p_1, \dots, p_n .

La probabilité d'un évènement E est alors donnée par : $P(E) = \sum_{i=1}^n P(E|r_i)$, où $P(E|r_i)$ est la probabilité de E dans le réseau r_i .

3.2. Utilisation dans le cadre de la modélisation de l'apprenant

3.2.1. Construction des différents réseaux

Nous nous limitons à l'analyse de réponses d'apprenant à des exercices dans des domaines où les connaissances de l'apprenant peuvent de décliner en savoirs et savoir-faire (typiquement, les mathématiques ou les sciences physiques).

Nous considérons donc trois types de nœuds : les *savoirs* et *savoir-faire*, qui constituent la couche domaine, et les *items* qui modélisent les réponses de l'apprenant et constituent quant à eux la couche tâche. Nous supposons de plus que nous disposons d'une représentation des connaissances du domaine sous forme de graphe non-orienté, dans lequel les *items* ne sont lié qu'à des *savoir-faire*.

L'interprétation traditionnelle causale des arcs dans un réseau bayésien (l'arc est orienté de A vers B lorsque A est une des causes de B) ne nous semble pas la plus pertinente dans le cas de la modélisation de l'apprenant. La question de savoir si un élève est bon en mathématiques parce qu'il résout beaucoup d'exercices ou au contraire qu'il résout beaucoup d'exercices parce qu'il est bon en mathématiques nous semble particulièrement difficile à trancher. Il en va de même pour ce qui est de discerner les éventuels liens de cause à effet entre la connaissance qu'a un apprenant des mathématiques et celles de ses objets. De ce fait, pour décider de l'orientation des arcs entre les différentes variables, nous avons recours à une analyse systématique *a priori* des dépendances entre les variables [HIBOU et al. 2006].

En résumé, cette analyse nous permet de confirmer la validité du choix habituellement fait dans la littérature d'une orientation des arcs des nœuds de la couche domaine vers ceux de la couche tâche, tout en confirmant l'impossibilité de déterminer *a priori* une orientation entre les nœuds *savoirs*.

3.2.2. Apprentissage des paramètres

Le cas d'utilisation des multi-réseaux que nous envisageons comporte une importante simplification par rapport au cas général : nous faisons l'hypothèse que les structures des différents réseaux considérés et les tables de probabilités conditionnelles sont connues. Seules sont à déterminer les probabilités *a priori* des différents réseaux et la loi de la variable cachée déterminant la probabilité de chaque réseau.

Dans la mesure où nous considérons que chaque apprenant est modélisé par un des réseaux, nous proposons d'utiliser l'algorithme EM [LAURITZEN 1995] après avoir affecté à chacun d'entre eux les données qui lui correspondent. Plus précisément (voir algorithme 1, les données sont notée d_1, \dots, d_n) après

initialisation au hasard de la probabilité de chacun des réseaux, on affecte à chaque donnée le réseau le plus probable, puis on apprend les paramètres manquants de chacun d'entre eux à l'aide de l'algorithme EM. On calcule ensuite la probabilité de chaque réseau sachant les données, puis on recommence la boucle, jusqu'à convergence des p_i .

3.2.3. *Diagnostics*

Un multi-réseau permet d'obtenir deux types de diagnostics : le diagnostic local et le diagnostic global.

Le diagnostic local s'obtient en calculant pour chaque variable de la couche domaine K la table de probabilités de K connaissant les valeurs des observables. Conceptuellement ce diagnostic est du même type que celui effectué dans un modèle de l'apprenant basé sur un seul réseau bayésien : il s'agit d'un calcul de probabilités conditionnelles. Cependant, ce calcul prend en compte les différents réseaux correspondants aux différents niveaux d'expertise.

Le diagnostic global est d'une nature tout à fait différente : il s'exprime non pas en terme de croyance sur l'acquisition de telle ou telle connaissance, mais en terme de réseau correspondant le mieux aux actions de l'apprenant. Ce réseau est celui qui maximise $P(r_i | d_j)$ parmi les réseaux pris en compte dans le modèle. Puisque les réseaux du modèle diffèrent en terme de structure, le diagnostic global donne des indications sur la structuration des connaissances de l'apprenant.

Il est important de remarquer que c'est la structure de multi-réseau qui donne un sens à l'expression « le réseau qui modélise le mieux l'apprenant ».

4. Expérimentations

Pour nos expérimentations, nous nous sommes servis de données recueillies lors de l'utilisation du logiciel de diagnostic en algèbre *PEPITE* [DELOZANNE *et al* 2003]. Nous ne cherchons pas à concevoir un modèle computationnel reprenant les éléments et les axes du diagnostic de *PEPITE*. Nous nous servons de *PEPITE* et des données utilisateurs qui ont été collectées durant plusieurs années comme champ d'expérimentation. Il s'agit pour nous de tester la validité de notre hypothèse sur le lien entre niveau d'expertise de l'apprenant et structure du réseau qui le modélise.

4.1. *Description du modèle*

Les exercices proposés dans *PEPITE* requièrent des connaissances en algèbre (sur les puissances, les racines carrées, les fractions,...) et en géométrie (sur les droites et les aires). Le squelette des couches domaines des différents réseaux bayésiens constituant le multi-réseau modèle de l'apprenant est ainsi constitué d'un nœud pour

chacun des *savoirs* algébriques ou géométrique nécessaires à la résolution des exercices (plus leurs déclinaisons en *savoir-faire*, comme par exemple *Calculer le coefficient directeur d'une droite*) auxquels s'ajoutent les nœuds *Maths*, *Géométrie* et *Algèbre*.

4.1.1. Les 5 structures retenues

En suivant notre analyse *a priori*, nous arrivons donc, à partir de ce squelette, à 8 réseaux différents (2 orientations possibles en chacun des trois nœuds *Maths*, *Géométrie* et *Algèbre*). Chacun de ces réseaux peut être interprété en terme d'homogénéité ou d'hétérogénéité des *savoirs* de l'apprenant (cf §2.2.2). Par exemple le réseau 6 représenté figure 4 modélise un apprenant aux connaissances homogènes en algèbre et hétérogènes en géométrie. En éliminant les réseaux qui ne s'interprètent pas de manière satisfaisante (par exemple un réseau modélisant un apprenant ayant des connaissances hétérogènes en algèbres mais homogènes en mathématiques), nous réduisons le nombre de composants du modèle de l'apprenant à 5 (figure 4).

La couche tâche des différents réseaux est constitué de 34 *Items* modélisant les réponses aux exercices, chacun d'entre eux étant connecté avec les *savoir-faire* à y mettre en œuvre.

4.1.2. Paramètres

Les relations entre les nœuds de la couche domaine ne sont pas bruitées, les nœuds racines (*i.e* ceux n'ayant pas de parents) suivant une loi équiprobable. Au niveau de la couche tâche, les paramètres représentent la probabilité de se tromper dans son domaine de compétence (uniformément fixé à 0,15), ou de deviner la réponse exacte (en fonction du nombre de réponses possibles).

La probabilité d'apparition de chacun des 5 réseaux est fixée à 0,2.

4.2. Résultats

Nous avons travaillé sur un ensemble de 400 fichiers de réponses aux exercices de *PEPITE*. Ces 400 apprenants sont des élèves scolarisés en classe de troisième ou de seconde. Pour chacun d'entre eux, nous utilisons la classe comme indicateur du niveau d'expertise, étant entendu que certains élèves de troisième peuvent être plus compétents que certains élèves de seconde.

Puisque le diagnostic global nous permet d'obtenir le réseau (et donc la structure) qui modélise le mieux l'apprenant, et que nous voulons tester une hypothèse sur le lien entre cette structure et le niveau d'expertise de l'apprenant, nous cherchons donc à voir s'il existe des différences significatives entre ces deux populations du point de vue du diagnostic global. Autrement dit nous voulons savoir s'il y a une corrélation entre le résultat du diagnostic global et le niveau scolaire.

Les différences de répartition obtenues entre les élèves de seconde et ceux de troisièmes ont été mesurées en utilisant le test du chi 2.

Nous avons fait des expérimentations avec ou sans apprentissage, celui-ci se faisant sur tout ou partie des données. Il faut noter que lorsque nous parlons d'apprentissage, il s'agit seulement de l'apprentissage des paramètres du multi-réseau et que l'indication de la classe de l'élève ne se trouve pas dans les fichiers qui servent à cet apprentissage.

Algorithme 1 :

Déterminer aléatoirement les p_i	
Répéter	
	Pour i variant de 1 à m et j de 1 à n calculer : $P(r_j d_i) = \frac{p_j P(d_i r_j)}{\sum_{k=1}^n p_k P(d_i r_k)}$
	Recalculer les p_j : $p_j = \frac{\sum_{i=1}^m P(r_j d_i)}{m}$
	Affecter chaque donnée d_i au réseau r_j tel que $j = \operatorname{argmax}_{k \in \{1, \dots, n\}} P(r_k d_i)$
	Apprendre les paramètres de chaque réseau à l'aide de l'algorithme EM à partir des données qui lui sont affectées.
Jusqu'à convergence des p_i .	

Nous donnons tableau 1 les résultats obtenus après apprentissage sur l'ensemble des données, l'ensemble des résultats que nous avons obtenus allant dans le même sens. Tout d'abord, l'hypothèse d'indépendance est rejetée (test significatif à 0,01). Ensuite, si nous nous intéressons à la répartition des différents réseaux dans les différentes classes, nous constatons qu'en seconde il y a sur-représentations des réseaux modélisant des apprenants aux connaissances homogènes en algèbre et sous-représentation de ceux modélisant des apprenants aux connaissances hétérogènes en algèbre, et que c'est l'exact opposé en classe de troisième.

Structure	Seconde				Troisième			
	Observé	Théorique	Écart		Observé	Théorique	Écart	
RB 8	104	118,15	1,69	-	109	94,85	2,11	+
RB 7	4	5,55	0,43	-	6	4,45	0,54	+
RB 6	25	24,41	0,01	+	19	19,59	0,02	-
RB 2	11	8,32	2,73	+	4	6,68	1,08	-
RB 1	69	56,58	5,73	+	33	45,42	7,14	-

Tableau 1. *Corrélation entre niveau scolaire et résultat du diagnostic global.*

Ces résultats confortent nos hypothèses. D'une part, nous avons bien mis en évidence une corrélation entre le niveau d'expertise et la structure du réseau modélisant l'apprenant, ce qui constitue un argument en faveur de la prise en compte de plusieurs réseaux au sein du même modèle.

D'autre part, nous constatons bien une sur-représentation des réseaux homogènes (ceux dont les arcs sont orientés du général vers le particulier) chez les apprenants les plus expérimentés et une sur-représentation des réseaux hétérogènes chez ceux qui le sont le moins.

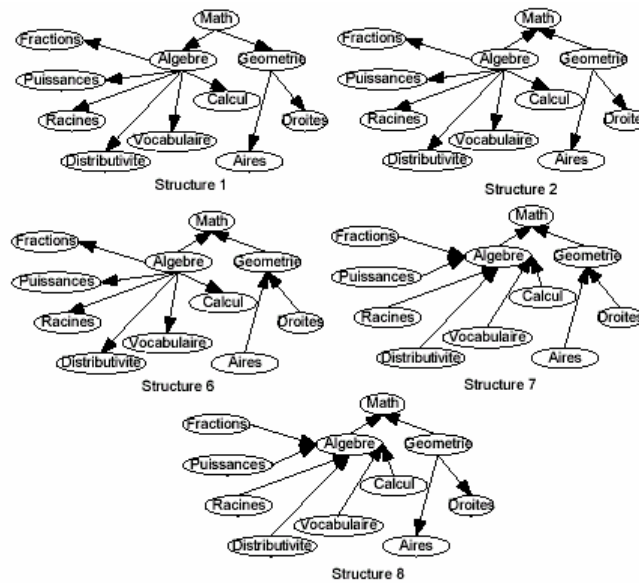


Figure 4. Les 5 réseaux bayésiens composant le modèle.

5. Conclusion

Nous avons montré en quoi d'un point de vue théorique et compte tenu de l'analyse de la littérature, il semble justifié de prendre en compte plusieurs réseaux bayésiens au sein d'un même modèle de l'apprenant, que se soit pour tenir compte des asymétries de dépendance ou de l'évolution structurelle des connaissances de l'apprenant. Nous proposons d'utiliser pour ce faire les multi-réseaux qui permettent une telle prise en compte dans un cadre mathématiquement bien fondé.

Les expérimentations présentées dans cet article sont autant d'arguments en la faveur de notre hypothèse sur le lien entre niveau d'expertise de l'apprenant et la structure du réseau bayésien qui le modélise.

Nous voyons principalement deux directions dans lesquelles poursuivre ce travail. D'une part en appliquant nos idées à des modèles conceptuels plus évolués que celui utilisé jusqu'à présent, et d'autre part en explorant l'exploitation qui peut être faite du diagnostic global, notamment en termes de remédiations.

Remerciements

Les auteurs tiennent à remercier Elisabeth Delozanne de nous avoir permis d'utiliser les données collectées lors de l'utilisation de *PEPITE*.

Bibliographie

- [CHI et al. 81] Chi M.T.H., Feltovitch P.J., Glaser R. (1981) : « Categorization and representation of physics problems by experts and novices », *Cognitive Science*, vol 5, pp 121-152.
- [CONATI et al. 02] Conati C., Gertner A., Vanlehn K. (2002) : « Using Bayesian networks to manage uncertainty in student modeling », *Journal of User Modeling and User-Adapted Interaction*, volume 12, n°4, pp 371-417.
- [DELOZANNE et al. 03] Delozanne E., Prévôt D., Grugeon B., Jacoboni P. (2003) : « Supporting teachers when diagnosing their students in algebra », Workshop Advanced Technologies for Mathematics Education, *supplementary Proceedings of Artificial Intelligence in Education*, Sydney, July 2003, IOS Press, pp 461-470.
- [GEIGER et al. 96] Geiger D., Heckermann D. (1996) : « Knowledge representation and inference in similarity networks and Bayesian multinets », *Artificial Intelligence*, volume 82 (1-2), pp 45-74.
- [HIBOU et al. 06] Hibou M., Labat J.-M. (2006) : « How to orientate arcs in a Bayesian network based student model? », *Proceedings of the 6th ICALT conference*, IEEE Conference Publishing, pp 560-562.
- [LAURITZEN 95] Lauritzen S.L. (1995) : « The EM-algorithm for graphical association models with missing data », *Computational Statistics and Data Analysis*, vol 1, pp 191-201.
- [MANSKE 05] Manske M., Conati C. (2005) : « Modelling Learning in an Educational Game », *Proceedings of the 12th Conference on Artificial Intelligence in Education*, C._K. Looi et al (eds), pp 411-418.
- [MAYO et al. 01] Mayo M., Mitrovic A. (2001) : « Optimising ITS behaviour with Bayesian networks and decision theory », *International Journal of Artificial Intelligence in Education*, n°12, pp 124-153.
- [MISLEVY 96] Mislevy R., Gitomer D. (1996) : « The role of probability-based inference in an intelligent tutoring system », *User-Modeling and User-Adapted Interaction*, n°5, pp 253-282.
- [PEARL 88] Pearl J. (1988) : *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, 1988.
- [REYE 04] Reye J. (2004) : « Student Modelling based on on belief networks, *International Journal of Artificial Intelligence in Education*, volume 14, pp 63-96.
- [VERMA et al. 91] Verma T., Pearl J. (1991) : « Equivalence and Synthesis of Causal Models », *Uncertainty in Artificial Intelligence*, vol 6, Elsevier Science Publishers, pp 220-227.
- [ZAPATA-RIVERA et al., 00] Zapata-Rivera J.-D., Greer J. : « Inspecting and Visualizing Distributed Bayesian Student Models », *Intelligent Tutoring Systems 00*, pages 544-553.